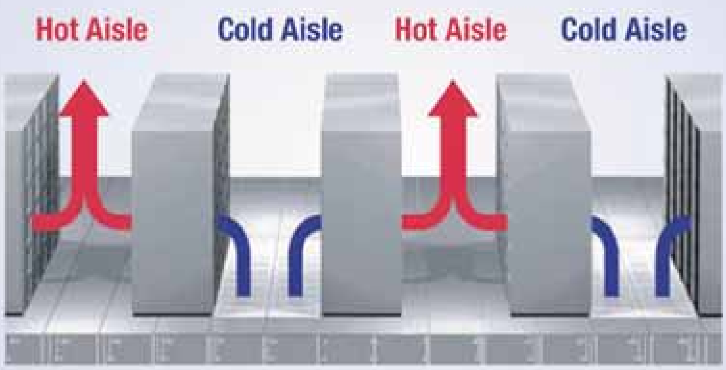
Big blades can be dangerous



Blade servers have experienced strong sales growth as more businesses look to save costs, space and ease the complexity of managing a host of individual servers. This article explores the impact upon power and cooling demands.

Data Centre Management, Dec 07

There are three drivers behind the seemingly ever increasing power demand per microprocessor: Number of transistors, the switching frequency and the device track spacing. If we compare the Intel Pentium IV to the groundbreaking devices of the early 70s we can clearly see the order of magnitude of technological and manufacturing innovation: An increase of over 4,000 times the number of transistors per chip to 42 million, whilst the power per chip increased from a few Watts in 1972 to the 40W of the Intel Pentium Pro of 1998 vintage and upwards to 130W/chip in 2006. Finally, in 2007-8 the power per chip will start to decrease but the number of chips per server may well increase to negate any advantage.

Why does the demand for processing capacity always stay ahead of the available technology? Despite server miniaturisation and increase in computing power by factors of tens of thousands, IT managers always want more computing capacity. It is never the case of taking out a row of five cabinets full of 'old' generation enterprise servers (15) and replacing them with less than half a cabinet full of 1U or a single blade array. The replacement will be planned to fill the same space (with 300 blades) and all of the newly available ICT applications will rapidly expand to fill the capacity and the floor power density could rise by 20:1.

Let's get physical

Regardless of what might be a good or bad solution, let us consider the physical limits that each of the main server options could create in a 42U high cabinet.

- **Enterprise Servers:** The previous generation big box server solution was 7-10U high and had dual main microprocessors with co-chips etc. The average power demand per server was around 700W, peaking at not much more than 1000W. In terms of density it was fairly common to see three such servers installed in one cabinet, with the physical upper limit of five a rarity. A typical cabinet full of enterprise servers was, therefore, in the region of 4kW.
- **Pizza Box Servers:** Mostly 1U form factor, each server draws around 250W on average and 350W peak. At 1U the theoretical population could be 42 in one cabinet - probably drawing in the region of 12kW.
- **Blade Servers:** Blade cassettes are available from many OEM's to house multiple blades in just a matter of a few U. Looking at two mainstream examples, the Dell PowerEdge 1855 (10 blades in 7U) and the IBM BladeCentre (14 blades in 8U), they consume around 550W/U.

Stuffing the box

These three server formats produce the possible power demand per cabinet from 4kW for enterprise servers, through 12kW for pizza box to a staggering 20kW for multiple blade cassette arrays. Note that a domestic oven is rated around 8-10kW.

So, what's the problem?

Consider the most common form of data centre floor space, the cold aisle/hot aisle arrangement: chilled air is fed from air con units around the room perimeter, under the raised floor and exiting upwards through perforated floor tiles into the cold aisles. The servers draw the chilled air through the front face of the cabinets and expel the hot waste air through the rear face into the hot aisle. The hot air is circulated back to the top of the down flow air con units and the waste heat extracted and dumped into the outside environment.

Space is always at a premium and the layout often allows no more than one full ventilated floor tile per cabinet footprint. So, in the worst case, the cooling air for one cabinet has to be delivered via one perforated floor tile unless the cabinets are not fully deployed or the aisles are deliberately widened etc. Now, there is a physical limit as to how much heat removal capacity can be delivered by a perforated floor tile. This is the eye of the needle: Tests have been done, rules of thumb used, computer models created and at the end of the day it all comes down to little more than 3kW of cooling capacity per 50% perforated tile that can be relied upon without using excessive air speed.

Expressing this in the more usual measurement of Watts/m^2 may be helpful: In a moderately dense cabinet layout a single cabinet has a gross footprint of $2m^2$ - equivalent to 1500W/m^2. This figure will resonate with data centre managers as a 'hot' room but when you consider the case of a particular cabinet being able to draw air from no more than two tiles it is easy to reach the conclusion that 5-6kW per cabinet is just about the practical limit for conventional cooling (and importantly for reliability).

The converse is also true: With most cabinets loaded above 5kW the conventional air cooling arrangements found in 99% of all data centres will not provide sufficient cooling and the load(s) will be subject to the risk of thermal alarm power shutdown.

What can we do?

In simple terms; find special solutions for the particular problem and treat hot spots in the data centre space. Until the server OEMs return to water/fluid cooling (like the first generation mainframes) we are stuck with air as the cooling medium.

Dedicated cooling elements (air to air and air to fluid heat exchangers) can be fitted to the cabinets, either in the doors or inside the base of the cabinet itself. These are available in the market place and have been proven up to 8kW. Overhead mounted cooling modules, capable of very much better targeted delivery of cool air down into the cold aisles and direct drawing up of the hot air (heat rises, cold sinks) push the cooling to 15kW per cabinet. Note that the data centre floor layout has to be planned well in advance for such solutions.

Some designs utilise a dialectric (non conducting) engineered fluid to remove the heat and so take away many of the objections of introducing water into the critical space and, in some solutions, above the critical load. A separate heat exchanger transfers the waste heat from the engineered fluid to the building's chilled water system. Single points of failure analysis should be carried out on all such cabinet dedicated internal and external solutions.

The two key problems in delivering sufficient cooling air into a single cabinet loaded above 12kW are linked, both the targeting and the redundancy:

- **Targeting:** The air has to be channelled both in delivery and return. Leakage and short circuiting of the air streams can no longer be tolerated.
- **Redundancy:** Any form of hot spot cabinet cooling that relies on a single cooling element or system, be that a door mounted heat exchanger or internally mounted fridge coil, has a fundamental flaw. Cooling shut downs, both planned for maintenance and unexpected failures, become as critical as the power supply. In the event of cooling failure the internal cabinet temperature will rapidly exceed the safe limits for the server hardware. This occurs in a matter of seconds rather than minutes. In a conventional solution multiple air con units feed the under floor plenum and the event of one unit failing (although it will cause a change in air flow pattern) will not lead to a catastrophic failure. Diverse power supplies and dual circuit chilled water feeds enable sophisticated redundancy and maintenance without shutdown schemes to be implemented.

With both these points in mind it is clear that twin cooling units, in N+N redundant format, feeding one or more cabinets with ducted air feed and ducted air returns, together with precise control of air volume, variable speed fan control and precise cooling capacity (step less compressor control) is the only practical and reliable solution.

Conclusions

In a conventionally designed data room do not allow the load in any individual cabinet to rise much above 5kW. At the same time keep the average load per cabinet to no more than 3kW, assuming the cooling system has the capacity (with redundancy) to remove that heat load. This results in a general rule: No more than one blade server cassette per cabinet unless special actions are taken.

If the density rises above 5kW in certain cabinets then special treatment of hot spots will be required. In this case the designer must take into account the criticality of the load and ensure that the cooling system has full redundancy and can cope with both instantaneous failure and maintenance without shutdown scenarios.

In the world of continuous computing and 24/7 uptime it is increasingly obvious that when the cabinet loads exceed an average of 5kW then the cooling system becomes as critical as the power system and UPS is required for both.

About the author

Ian F Bitterlin
BSc(Hons) DipDesInn MCIBSE MIET

With over 36 years experience of writing technical articles for leading companies and institutions, Ian is a world renowned author and speaker and an expert in all aspects of critical power and building services.